SIM White Light On–board
Processing Algorithms

Mark H. Milman and Joe Catanzarite

Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109

1. Introduction Interferometry in optical astronomy is an important and growing field of astronomical observation. A number of stellar interferometers have come online over the past several years (e.g. Palomar [1], NPOI [2], REGAIN [3]), and several more are due to be operational in the near future (e.g. Keck [4], VLTI [5]). In addition, space based interferometers are also planned missions of NASA's Origins program [6], including the Space Interferometry Mission (SIM), the focus of the present paper. The fundamental measurement made by each of these interferometers is the white light fringe measurement to determine the optical pathlength delay between the two arms of the interferometer. SIM makes white light measurements with three independent interferometers observing three different objects. Two of these are the "guide" interferometers that observe bright objects (approximately 7th magnitude) to track the rigid body motion of the instrument. The third interferometer, the "science" interferometer, observes the science targets of interest.

A common method for making the white light measurement is to disperse the interfered starlight across multiple spectral channels and employ a phase shifting interferometry (PSI) algorithm to determine the phase difference in each channel. In the phase–delay method, knowledge of the mean wavenumber of each channel is used to convert the phase measurement to a delay measurement. The individual delays are then combined to produce a single delay. There are many alternatives for performing this last function – the delays can simply be averaged, or they can be combined in a weighted least squares sense, etc. When this delay measurement is combined with a metrology measurement of the pathlength difference of the starlight as it traverses the optical system, the fundamental observable of the instrument is constructed: the "external" pathlength delay, which is mathematically equivalent to the projection of the interferometer baseline vector onto the star direction vector. From these measurements the astrometric parameters of the star can be estimated.

This paper focuses on several of the important aspects of how this fundamental observable is constructed, and some of the errors that arise in its development. There is a significant history of PSI algorithms that deal with commonly associated errors in interferometry measurements. These include problems associated with small signal, changes in the pathlength (e.g. vibration) as the measurement is being made, non–monochromatic light with error in the mean wavelength in spectral channels, etc. In many instances algorithms can be developed to overcome these errrors ([7],[8], [9]) Here we will tie together the interferometric measurement with the metrology measurement especially to overcome the effects of vibration, and errors inherent in the modulating element. Also some preliminary investigation into the impact of non–monochromatic light has PSI algorithms will be made.

2. Background. SIM makes the pathlength delay measurement by a combination of internal metrology measurements to determine the distance the starlight travels through the two arms of the interferometer, and a measurement of the white light stellar fringe to find the point of equal pathlength. Figure 1 defines the basic geometry of the interferometer. Assume here that the optical axes of the two telescopes that comprise the interferometer apertures are aligned with the star position vector s. The planes P_X and P_Y are two planes of equal phase for the planar wavefront of the starlight, and each is normal to s. The light through the two arms combine at the beam combiner located at z. Let l_X and l_Y denote the internal optical pathlength through the X and Y arms, respectively, to z. And let T_X and T_Y denote the total pathlengths of the starlight through the X and Y arms. Define

$$E_X = T_X - l_X; \quad E_Y = T_Y - l_Y. \tag{1}$$

Then $E_Y - E_X$ is the external pathlength, and defines the basic astrometric equation which relates the star direction vector and the interferometer baseline vector,

$$E_Y - E_X = \langle s, Y - X \rangle, \tag{2}$$

where $Y - X$ is the baseline vector of the interferometer. Now $E_Y - E_X$ is not directly observable, but from the equation (deduced from (1))

$$E_Y - E_X = (T_Y - T_X) - (l_Y - l_X), \tag{3}$$

we see that each of the quantities on the right is observable. $T_Y - T_X$, the total pathlength difference, is measured by means of white light fringe estimation, and $l_Y - l_X$, the internal pathlength difference is measured by a metrology system employed by the instrument.

Eq. (2) is typically written as

$$d_{ext}(t) = \langle b(t), s \rangle + \nu(t) \tag{4}$$

where d_{ext} is the instantaneous external pathlength delay, $b(t)$ is the instantaneous interferometer baseline vector, and ν is the measurement error. The fundamental objective of the instrument is to make measurements so that the astrometric parameters of the star vector s can be determined. Additional "guide" interferometers and an external metrology system that ties together all of the interferometers are used to track the baseline vector $b(t)$. Because a finite integration time is required to make this measurement, the basic model for the astrometric equation has the form

$$\bar{d}_{ext} = \langle s, \bar{b} \rangle + \nu, \tag{5}$$

where the overbar represents a time–averaged quantity. The emphasis of the remainder of the paper is the synthesis of $\bar{d}_{ext}$ via time–averaging of the white light fringe estimates and the internal metrology measurements over the observation period. Obtaining $\bar{b}$ involves a similar process using the guide interferometers and auxillary metrology measurements to relate the three interferometer baseline vectors. This process, sometimes refered to as "baseline regularization" [*] will not be discussed here.

3. Measuring the external delay. As described in the previous section, the external delay measurement, d_{ext} is synthesized from two other delay measurements – the total delay and the

internal delay. Henceforth these latter two delays will be denoted d_{tot} and d_{int}. respectively. The three different delays are related:

$$d_{tot} = d_{ext} + d_{int}.$$ (6)

The initial focus is on how to compute the average external pathlength delay for a single white light fringe measurement that requires, say τ sec of integration time. We will assume that this integration time coincides with the modulation period for the phase shifting interferometry (PSI) measurement. A nominal value for τ we will use corresponding to a typical science target is 100ms; although dim targets may require several seconds of integration time. We will first take a small digression to explain how d_{tot} is ideally measured using white light interferometry

A general perspective of building phase estimators uses the following simple idea. We start with the fundamental interferometric intensity equation for monochromatic light

$$y = I_0[1 + V(\cos(kx + \phi))],$$ (7)

where y is the observed intensity, I_0 is the dc intensity, V denotes the visibility, k is the wavenumber of the monochromatic light (or mean wavenumber over a spectral channel), x is a known modulator displacement, and ϕ is the sought after phase. Phase shifting interferometry entails introducing known pathlength changes via the x variable in the expression above to set up a nonlinear system of equations to solve for all of the unknown variables. The common way for solving this system is to introduce the state X consisting of three components, $X \equiv (I_0, I_0 V \cos(\phi), I_0 V \sin(\phi))$. It can be shown [7] that the system of equations is *linear* in X,

$$Y = AX,$$ (8)

where Y is the vector of observed photon counts at different values of pathlength change variable x, and A is a known matrix that maps the state vector into the observation vector.

An N–bin algorithm essentially uses N values of x in (7) to generate N values of the observed intensity y. Hence, N equations in the three unknown variables. For an $N-$ bin "integrating bucket" algorithm (see [12]) centered at zero we define s as the total stroke length of the phase modulator and we let λ denote the wavelength of the light. From these two variables we define $\gamma = s/\lambda$ and $\Delta = 2\pi\gamma/N$. In the integrating bucket method, the modulator linearly sweeps through $2\pi\gamma$ radians from $-\pi\gamma$ to $\pi\gamma$. The design matrix A for the integrating bucket method that relates photon counts to the state is obtained by integrating (7) over the "buckets" $(u_i - \Delta/2, u_i + \Delta/2)$ with respect to x. The resulting A matrix has the form (cf [12])

$$A = \begin{bmatrix} \Delta & \cos(u_1) & -\sin(u_1) \\ \vdots & \vdots & \vdots \\ \Delta & \cos(u_N) & -\sin(u_N) \end{bmatrix},$$ (9)

where $u_i = (i - 1/2 - N/2)\Delta$. The estimate of X is obtained from *any* unbiased linear estimator of the state, and the phase is subsequently derived from X via $\phi = \arctan(X_3/X_2)$. Once the phase has been determined, the delay is calculated by using the wavelength of the light. To avoid multiple wavelength ambiguities in the delay, the white light is dispersed into multiple spectral channels, each with a different mean wavelength. The ambiguities can be resolved by using the multiple wavelengths [14]. A question that is addressed in Section 5 is the validity of using monochromatic algorithms in a spectral channel that has a nonvanishing width.

3

So keping in mind the need to modulate the pathlength in order to measure it, we write

$$d_{tot}(t) = d_{ext} + d_{int} + d_{pzt}(t), \qquad (10)$$

where $d_{pzt}(t)$ is the pathlength introduced by the modulation. If everything is "perfect", the white light fringe estimate is an estimate of $d_{ext} + d_{int}$, which presumes that this value is constant over the τ second integration time. (In reality this is not the case, and this important non–ideality will be discussed in the next section.) Call this measurement y_{wl};

$$y_{wl} = d_{ext} + d_{int}. \qquad (11)$$

Keeping in mind that the internal pathlength is changing due to the phase modulation, the measurement made by internal metrology has the form

$$m(t) = d_{int} + d_{pzt}(t). \qquad (12)$$

The average value of m is d_{int} since $\bar{d}_{pzt} = 0$, thus

$$d_{ext} = y_{wl} - \bar{m}. \qquad (13)$$

4. Mechanical white light error sources and some fixes.

There are many different sources of errors contributing to white light fringe estimation. In this section we will focus on the class that is generated by changes in the pathlength while the phase is being measured. When the pathlength is changing, even if we interpret eq.(13) as the average external pathlength difference over the τ second integration period (which is indeed the quantity that is required for the astrometric relationship in (2)), there in general will be an error because y_{wl} may not compute the average total pathlength delay over the integration time.

One immediate way of generating an error of this type is if the motion of the modulator deviates from the motion assumed in the phase estimation algorithm, i.e. the matrix A in (9) which is constructed from an ideal modulator motion is in error. A straightforward way around this problem is to use the internal metrology measurement to create a new matrix A with every modulation stroke. Unfortunately this is not entirely desirable because in addition to building this matrix for each phase measurement, its pseudoinverse also has to be constructed to generate the phase delay. An alternative to this is to use an algorithm based on an assumed and fixed trajectory of the modulating element, and then make corrections to the computed phase based on measured values obtain from internal metrology. The main advantage here is simplicity, with payoffs including a reduced computational load, and very importantly, a simpler analysis of errors.

Following this tack, denote the modeled trajectory by d_{pzt}^{mod} and write

$$d_{pzt}(t) = d_{pzt}^{mod}(t) + [d_{pzt}(t) - d_{pzt}^{mod}(t)]. \qquad (14)$$

Upon defining

$$d_{int}^* = d_{int} + [d_{pzt}(t) - d_{pzt}^{mod}(t)], \qquad (15)$$

(10) becomes

$$d_{tot}(t) = d_{ext} + d_{int}^*(t) + d_{pzt}^{mod}(t). \qquad (16)$$

So now

$$y_{wl} = d_{ext} + \bar{d}_{int}^*, \tag{17}$$

with metrology measurements

$$m(t) = d_{int}^* + d_{pzt}^{mod}(t). \tag{18}$$

Averaging the metrology measurements yields

$$\bar{m} = \bar{d}_{int}^*, \tag{19}$$

since $\bar{d}_{int}^{mod} = 0$. Thus

$$y_{wl} - \bar{m} = \bar{d}_{ext}, \tag{20}$$

so that

$$\bar{d}_{ext} = y_{wl} - \bar{m}. \tag{21}$$

Although we have apparently circumvented the need for redefining an A matrix with each stroke, the probelm is not completely solved. We have introduce the time varying term $d_{pzt}(t) - d_{pzt}^{mod}(t)$ into the total path.

2. White light systematic errors. Define $\delta_{ext}(t)$ and $\delta_{int}^*(t)$ by

$$\delta_{ext} = d_{ext}(t) - \bar{d}_{ext}, \quad \delta_{int}^*(t) = d_{int}^*(t) - \bar{d}_{int}^* \tag{22}$$

and set

$$\delta \equiv \delta_{ext} + \delta_{int}^*. \tag{23}$$

Suppose an $N-$bin integrating bucket algorithm is used to convert the vector of photon counts $(y_1, ..., y_N)$ into phasor estimates $I_0 V \cos(\phi), I_0 V \sin(\phi)$ by the $3 \times N$ gain matrix $K = (k_{ij})$,

$$I_0 V \cos(\phi) = \sum_{j=1}^{N} k_{2j} y_j, \quad I_0 V \sin(\phi) = \sum_{j=1}^{N} k_{3j} y_j, \tag{24}$$

where $I_0 =$ intensity, $V =$ visibility and $\phi =$ phase. The true average delay over the single phase measurement period is $\bar{d}_{ext} + \bar{d}_{int}^*$ with associated phase ψ,

$$\psi = \frac{2\pi}{\lambda}(\bar{d}_{ext} + \bar{d}_{int}^*), \tag{25}$$

where λ denotes the wavelength of the light. In [9] it is shown that the error in the estimate using the phase derived from (24) is

$$\psi - \phi = \frac{1}{2\sin(\Delta/2)}\{\sum k_{3j}[\cos^2(\psi) \int_{u_i-\Delta/2}^{u_i+\Delta/2} \sin(u)\delta(u/2\pi\gamma)du + \sin(2\psi)/2 \int_{u_i-\Delta/2}^{u_i+\Delta/2} \cos(u)\delta(u/2\pi\gamma)du] +$$

$$\sum k_{2j}[\sin(2\psi)/2 \int_{u_i-\Delta/2}^{u_i+\Delta/2} \sin(u)\delta(u/2\pi\gamma)du + \sin^2(\psi) \int_{u_i-\Delta/2}^{u_i+\Delta/2} \cos(u)\delta(u/2\pi\gamma)du]\} \tag{26}$$

Note that in general the white light error due to the pathlength variation δ is also a function of the offset ψ. For SIM applications ψ is typically small (a fraction of a radian), and the first term dominates the error.

A frequency response of the error can be generated for any given gain matrix K by fixing a phase offset for ψ and computing the resulting error for deviations of the form $\delta(t) = \sin(\omega t) - \mu(\omega)$, where $\omega = 1$ corresponds to the modulation frequency and $\mu(\omega)$ is the average value of $\sin(\omega t)$ over the modulation interval. Because the astrometric observable entails an average value of the external pathlength delay over several seconds of integration time on a typical science object, an important quantity is the average value of the error due to sinusoid variation of frequency ω over many, say M, modulation periods. (For example, if a 30sec integration time is required and a delay measurement is made every .1sec, then $M = 300$.) In [10] this error is shown to have a form (ignoring the variations in ψ from measurement to measurement)

$$E_{ave}(\omega) = f(\omega)C_M(\omega) + g(\omega)S_M(\omega), \tag{27}$$

where

$$C_M(\omega) = \frac{1}{M}\sum_{k=0}^{M-1}\cos(2\pi k\omega), \quad S_M(\omega) = \frac{1}{M}\sum_{k=0}^{M-1}\sin(2\pi k\omega), \tag{28}$$

and the functions f and g decrease as $1/\omega$ with increasing frequency. The sums above can be computed analytically to obtain (using the complex exponential forms for sin and cos and recognizing the sums as a geometric series),

$$S_M(\omega) = \frac{\sin(2(M-1)\pi\omega) + \sin(2\pi\omega) - \sin(2M\pi\omega)}{2M(1-\cos(2\pi\omega))}, \tag{29}$$

and

$$C_M(\omega) = \frac{1-\cos(2\pi\omega) + \cos(2\pi(M-1)\omega) - \cos(2M\pi\omega)}{2M(1-\cos(2\pi\omega))}. \tag{30}$$

Thus when ω is an integer, there is no attenuation in the error due to averaging. But when it is not, the error will decrease as $1/M$ for each frequency ω. Because of the denominator terms, it is seen that the attenuation is slower when the frequency is close to an integer. The oscillation in the error will also be slower for frequency values near an integer. This oscillation frequency is $\omega_0 \equiv \omega \bmod 1$, so that with M measurements there would be $M\omega_0$ total oscillations. In contrast to the case of a periodic disturbance with an integer frequency, the expectation for an arbitrary signal based on Fourier analysis is that the average error diminishes with time (increasing number of intervals). Thus, periodic disturbances that are close to an integral multiple of the modulator frequency will attenuate slowly (and not at all if the disturbance is an exact multiple).

The error formula (26) can be used to correct the phase error if the deviation function δ is known. We will begin by assuming this is the case, and then discuss how this can be implemented and identify the residual errors.

Simplifying the notation, we may write (26) as

$$\psi - \phi = \int G(u,\psi)\delta(u)du, \tag{31}$$

where $G(u,\psi)$ is formed directly from the gain definition and the true phase ψ. Suppose $\delta(u)$ is known, and write $e = \psi - \phi$. Then (31) can be written as

$$e = \int G(u,\phi)\delta(u)du + \int \frac{\partial G(u,\phi)}{\partial \phi}e\delta(u)du. \tag{32}$$

6

Then,

$$e = \int G(u,\phi)\delta(u)du\left(1 - \int \frac{\partial G(u,\phi)}{\partial \phi}\delta(u)du\right). \tag{33}$$

As $\delta(u)$ is presumably small, the second term in the expression above should be ignorable, and we are left with

$$\psi = \phi + \int G(u,\phi)\delta(u)du. \tag{34}$$

The implementation of the correction term in (34) is straightforward. Let (k_{ij}) denote the $3 \times N$ components of the gain matrix. Then for $i = 1, ..., N$ the following quadratures are computed:

$$s_i = \int_{u_i-\Delta/2}^{u_i+\Delta/2} \sin(u)\delta(u/2\pi\gamma)du, \qquad c_i = \int_{u_i-\Delta/2}^{u_i+\Delta/2} \cos(u)\delta(u/2\pi\gamma)du. \tag{35}$$

Then (36) is implemented as

$$\psi = \phi + \frac{1}{2\sin(\Delta/2)}\left\{\cos^2(\psi)\sum_{i=1}^{N} k_{3i}s_i + \frac{\sin(2\psi)}{2}\sum_{i=1}^{N} k_{3i}c_i + \frac{\sin(2\psi)}{2}\sum_{i=1}^{N} k_{2i}s_i + \sin^2(\psi)\sum_{i=1}^{N} k_{2i}c_i\right\}. \tag{36}$$

Recall that δ is the deviation about the mean of the sum of the external path and the modified internal path. This quantity is actually somewhat observable for the science interferometer.

4. Sources of error. The mechanical sources for error in the determination of the white light estimate are those mechanisms that directly contribute variations about the mean of the external path and variations about the mean of the modified internal path defined in (1) and (15), respectively.

Since the star direction vector may assumed to be fixed in inertial space over the duration of the observation, the only way the external path can change is by a change in the inertial position of the interferometer baseline vector. This vector is defined as the difference between the two fiducial positions on the siderostat mirrors, and is thus affected by both rigid body motion of the instrument and vibrations propagating through the siderostats. The only means for active control of the external pathlength is via the attitude control system. Because the bandwidth of the attitude control system is between .1Hz and .01Hz, and the modulation period is typically less than a second (.1sec for grid star, faster for a guide star), the rigid body contribution to δ^*_{ext} is the dominant term for the science interferometer and is essentially linear. (A representative disturbance spectrum for the felxible body contribution is being developed.) The pathlength control system compensates for the rigid body motion to maintain the total pathlength error to the 10nm requirement to ensure good fringe visibility. The controller therefore introduces a large change in the internal path for the science interferometer. Assuming the rigid body motion of the instrument is a sinusoid with amplitude of 10^{-5}rad (about 2asec) and a frequency of .01Hz, the maximum external pathlength velocity is about 6μm/sec. If a phase measurement is made every .1sec, as is assumed for the science interferometer, the pathlength controller must vary the internal pathlength by .6μm to compensate for the change in the external pathlength over the phase measurement time period to achieve the 10nm requirement. Hence, the presence of pathlength control mitigates the external and internal pathlength variations that contribute to the white light estimation error. In fact the guide interferometers and external metrology provide the signal that drives the compensation and if the control system acts ideally, the internal pathlength changes introduced by the controller completely cancel the external pathlength changes.

In the real situation the compensation is effective over some specified bandwidth, and a residual variation remains with spectral content beyond the bandwidth of the controller. Also what remains uncompensated are the inaccuracies of the signal, the implementation of the controller, internal vibrations/drifts, and modulator motion error. Using a fixed gain phase measurement algorithm, the contribution of modulator motion to the error is the difference between the assumed motion and the true motion of the modulator.

5. Non–monochromatic problem. The analysis thus far has centered around the monochromatic light case. Modifications are necessary when the spectral bandwidth is not vanishingly small. In this section we will take a preliminary look at some of the parameters that may mandate these modifications. We examine three parameters: the number of spectral channels (which determines the channel width), the number of time bins, and the pathlength delay that is to be estimated. For example, increasing the number of channels eventually reduces to the monochromatic case. So one question is how many channels are necessary to make the problem look monochromatic? And when this is not possible, what modifications of the baseline monochromatic estimator are necessary to recover performance? We will be treating only the simplest model that assumes a rectangular bandpass with constant intensity/visibility within the channel. For simplicity we will only analyze the situation using a phase stepping (as opposed to integrating bucket) modulation scheme.

When the monochromatic estimator is not sufficient for this model there are essentially two modifications that can be made. The simplest one involves including a known sinc function in the matrix equation that relates the state variables (intensity and phasor quantities) to the measured intensities. The more complex model contains the unknown delay as part of the sinc term. This model leads to a nonlinear estimation problem to determine the state variables. We will focus attention throughout this discussion on a single spectral channel with mean wave number k_0 ($k_0 = 2\pi/\lambda_0$, with λ_0 denoting the corresponding wavelength) and channel width Δk. We will assume throughout that $\Delta k = 2\pi/(10^3 M)$ (units of 1/nm) where M denotes the number of spectral channels. (This value of Δk approximately corresponds to a spectral band between 500nm and 1000nm divided into M channels.) For a pathlength difference, x, the intensity model using a rectangular bandpass is [1]

$$I = I_0[1 + V\mathrm{sinc}(\Delta k x/2)\cos(k_0 x)]. \tag{37}$$

If δ denotes the unknown delay, and the pathlength is modulated in N equidistant steps, $x_i = i\lambda_0/N$, $i = 1, ..., N$, the estimation problem is to determine δ from the N equations

$$y_i = I_0[1 + V\mathrm{sinc}(\Delta k x_i/2 + \Delta k \delta/2)\cos(k_0 x_i + k_0\delta)] \tag{38}$$

In what follows y_i, k_0, Δk, and x_i are all assumed to be known, i.e., we are postulating a phase stepping method in which there is no measurement error, the wavenumbers and spectral channels widths have been precisely calibrated, and the phase steps have no error.

It is useful to relate the value of the sinc function in (1) to unity to capture the comparison between the models (2) and (3). If M channels are used and the modulation sweeps a micron, the maximum deviation of $\mathrm{sinc}(\Delta k x/2)$ from unity is about $\pi^2/(6M^2)$; which provides some measure of the deviation between the two models. To truly quantify the errors in the estimate of the delay it is necessary to solve for the delays in the two models.

The ususal manipulation of (3) leads to the linear estimation problem from the system

$$
\begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} 1 & \cos(u_1) & -\sin(u_1) \\ \vdots & \vdots & \vdots \\ 1 & \cos(u_N) & -\sin(u_N) \end{bmatrix} \begin{bmatrix} I_0 \\ I_0\cos(\phi) \\ I_0\sin(\phi) \end{bmatrix}, \tag{39}
$$

where $u_i = k_0 x_i$ and $\phi = k_0\delta$. (The slightly different form of the equations is due to the phase–stepping versus integrating–bucket modes of phase modulation.) We will write the nominal monochromatic system of equations as

$$
y = A_0 x. \tag{40}
$$

In the figure below we plot the true delay versus the error in the delay estimate based on the model (3). These errors were computed for four different combinations of time bins and spectral channels. The wavelength and modulation length are both 750 nm in these plots.
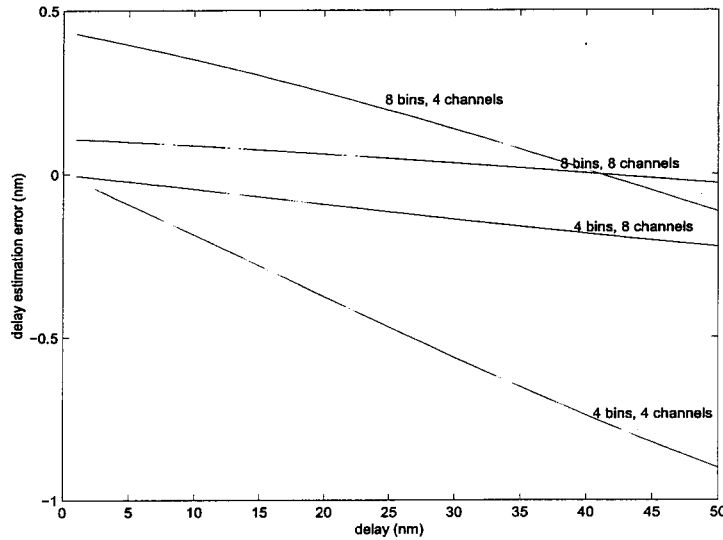

Figure 1. Phase estimation error using monochromatic alogrithm

Thus it is seen that using eight spectral channels produces a relatively small error (although not necessarily acceptable for all applications). This error is what is obtainable without modifying the basic algorithm.

In the next figure the number of channels is increased to 16, and two modulation lengths were used: 750 nm and 1000 nm, while the wavelength was fixed at 750 nm. The resulting error in the worst case is less than 50pm.
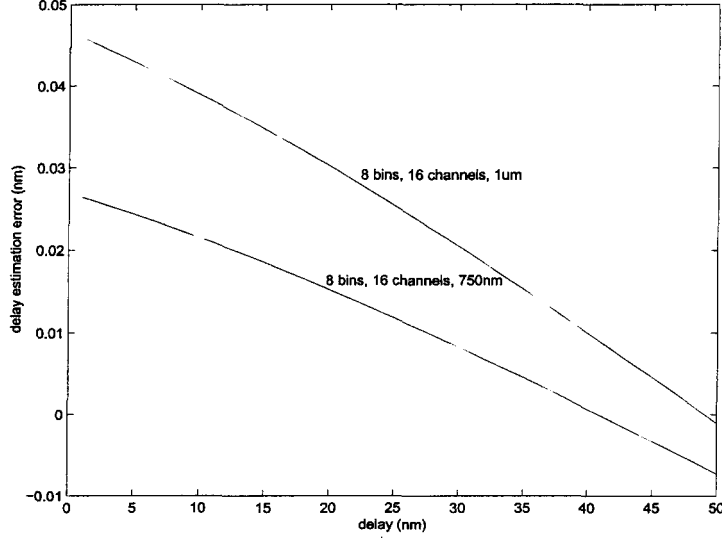

Figure 2. Phase estimation error using monochromatic alogrithm with 16 spectral channels

Next we model (1) using the sinc function, but only at the known modulated delays. This model has the form

$$
\begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} 1 & \mathrm{sinc}(\Delta k x_1/2)\cos(u_1) & -\mathrm{sinc}(\Delta k x_1/2)\sin(u_1) \\ \vdots & \vdots & \vdots \\ 1 & \mathrm{sinc}(\Delta k x_N/2)\cos(u_N) & -\mathrm{sinc}(\Delta k x_N/2)\sin(u_N) \end{bmatrix} \begin{bmatrix} I_0 \\ I_0\cos(\phi) \\ I_0\sin(\phi) \end{bmatrix}, \tag{41}
$$

which we write in matrix form as

$$
y = Ax. \tag{42}
$$

In anticipation of a larger nonlinearity that results when the number of spectral channels is decreased and/or the delay offset becomes larger, we introduce the matrix function $A(\delta)$,

$$
\begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} 1 & \mathrm{sinc}(\Delta k[x_1+\delta]/2)\cos(u_1) & -\mathrm{sinc}(\Delta k[x_1+\delta]/2)\sin(u_1) \\ \vdots & \vdots & \vdots \\ 1 & \mathrm{sinc}(\Delta k[x_N+\delta]/2)\cos(u_N) & -\mathrm{sinc}(\Delta k[x_N+\delta]/2)\sin(u_N) \end{bmatrix} \begin{bmatrix} I_0 \\ I_0\cos(\phi) \\ I_0\sin(\phi) \end{bmatrix}. \tag{43}
$$

With this notation, $A = A(0)$.

The phase estimate is recovered exactly as before, viz. $\hat{\phi} = \arctan(\hat{x}_3/\hat{x}_2)$ where $\hat{x} = A^{\dagger}y$. ($A^{\dagger}$ is the pseudoinverse of A.) It is important to realize that this is still a linear estimation problem. Figure 3 contains results, analogous to those in Figure 1. It is seen that an eight bin/eight spectral channel algorithm yields a very small error that should suffice for almost all applications. However,

the use of eight spectral channels may increase the read noise penalty to be unusable on dim stars. In the worst case (four bins, four channels, 50nm offset), an error of about 85 pm results.

Because the results are somewhat insensitive to the small error in the model of the sinc function (using $A(0)$ instead of $A(\delta)$), we conjecture that error analysis of this algorithm does not have to include the effects of (small) errors introduced through the sinc term in (1). Hence, much of the analysis should very closely follow the analysis that has been worked through for monochromatic light.
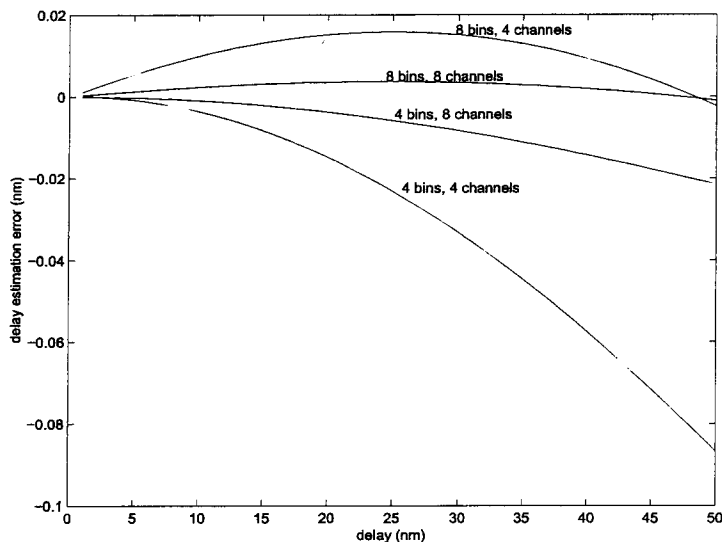

Figure 3. Phase estimation errors due to imprecise modeling for 4 configurations.

5.1. A nonlinear estimation scheme. For larger values of Δk and δ (or more stringent requirements) a nonlinear approach to the estimation problem is necessary. Thus we introduce the function

$$G(x) = [A(\delta(x))]^\dagger y, \tag{44}$$

where $\delta(x) = \tan^{-1}(x_3/x_2)$, and determine a fixed point of the map. That is we solve the equation

$$x = G(x). \tag{45}$$

This equation is solved via the fixed–point teration scheme

$$x_{k+1} = G(x_k), \quad x_0 = G(0) = A(0)^\dagger y. \tag{46}$$

Convergence of the scheme is motivated as follows. First we note that a fixed point exists. (This is guaranteed by (2). In the case of noise on the measurements, a regularity argument must be used to prove existence.) Given that a fixed point exists, local convergence of the iteration (11) is established if it can be shown that $|G'(x)| < 1$ in a neighborhood of the solution. An important approximation leading to this result is (see [3])

$$|A^\dagger - B^\dagger| \le |A^\dagger||A - B|_F, \tag{47}$$

11

where $|\cdot|_F$ denotes the Frobenius norm. Thus we are led to an analysis of $d/dx[A(\delta x)y]$, and to find conditions so that this the norm of this is small. But the components of the matrix A containing δx are of the $\mathrm{sinc}(r + \delta x)$ where $\delta x = \tan^{-1}(x_3/x_2)$. The sinc function is relatively flat for small values of r, and by operating with δx near zero, the derivatives are small. These arguments (made carefully) lead to the bound $|G'| < 1$ in a neighborhood of the solution, and thereby justifying the iteration scheme.

An application of this iteration that may be of some interest is to consider the solution to the problem with a single spectral channel. This wide spectral band leads to a large value of Δk, and the phasor estimation problem becomes nonlinear. In Figure 4 we plot the error in the delay estimate using the linear model (6) (with wavelength=900nm).
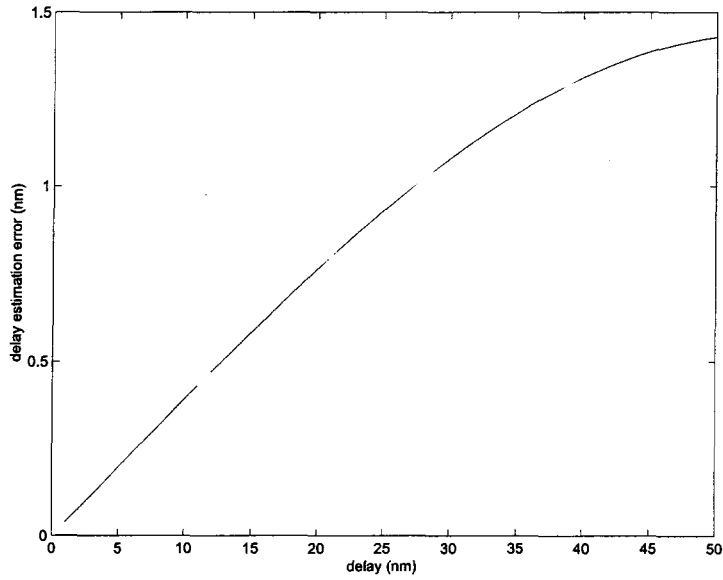

Figure 4. Phase estimation errors due to linear model.

12

Note that over a nanometer of error results. Next we invoke the iteration scheme (11). The errors from the scheme are shown in Figure 5 for one and two iterations. The maximum error after two iterations is less than 2 pm. A third iteration leads to a maximum error that is sub–picometer.
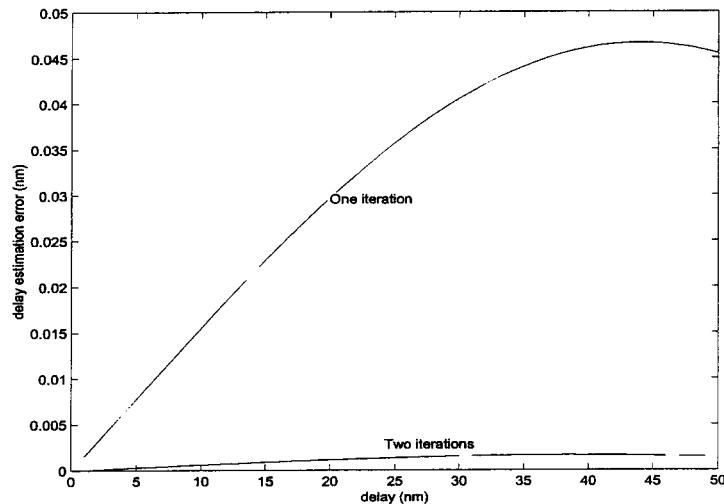

Figure 5. Phase estimation error using iteration scheme.

The idea of using a single spectral channel is a little intriguing from an SNR perspective, but is doubtful whether this simplified model of such a wide spectral channel is a faithful representation of the physics. For example, we probably cannot assume that the light has uniform intensity across the band, and this in turn affects the definition of the mean wavelength of the radiation [2]. But getting a handle on these sensitivities is worthwhile in general, since there will be a trade in performance between the number of channels, SNR, and the magnitude of errors due to wavelength dependent phenomena.

Acknowledgement

References

[1] M. M. Colavita, J. K. Wallace, B. E. Hines, Y. Gursel, F. Malbet, D. L. Palmer, X. P. Pan, M. Shao, J. W. Yu, A. F. Boden, P. J. Dumont, J. Gubler, C. D. Koresko, S. R. Kulkarni, B. F. Lane, D. W. Mobley, and G. T. van Belle, "The Palomar Testbed Interferometer," The Astrophysical Journal, 510, pp. 505–521, January, 1999.

[2] C. A. Hummel, "Practice of interferometry with NPOI," in *Optical Interferometry in Optical Astronomy*, P. J. Lena and A. Quirrenbach, Ed., Proc. of SPIE, Vol. 4006, 2000, pp. 459–464.

13

[3] D. Mourard, D. Bonneau, A. Glentzlin, G. Merlin, R. Pétrov, M. Pierron, N. Thureau, L. Abe, P. Bério, A. Blazit, O. Chesneau, P. Stee, S. Ragland, F. Vakili, and C. Verinaud, "The G12T/REGAIN Interferometer," Ibid, pp. 434–443.

[4] M. M. Colavita and P.L. Wizinowich, "Keck Interferometer: Progress report," Ibid, pp. 310–320.

[5] A. Glindemann, R. Abuter, F. Carbognani, F. Delplancke, F. Derie, A. Gennai, P. Gitton, P. Kervella, B. Koehler, S. Lévêque, S. Menardi, A. Michel, F. Paresce, T. P. Duc, A. Richichi, M. Schöller, M. Tarenghi, A. Wallander, and R. Wilhelm, "The VLTI interferometer: A unique instrument for high–resolution astronomy," Ibid, pp. 2–12.

[6] R. Danner and S. Unwin, *Space Interferometry Mission, Taking the Measure of the Universe*, JPL Pub. 400–811, March, 1999.

[7] K. Creath, "Phase–measurement interferometery techniques," in *Progress in Optics XXVI*, Elsevier Science Pub. B. V., 1988, pp. 350–393.

[8] P. J. de Groot, "Derivation of algorithms for phase–shifting interferometry using the concept of a data–sampling window," Appl. Opt., Vol. 34, August, 1995, pp. 4723–4730.

[9] J. Schwider, R. Burrow, K.–E. Elsner, J. Grzana, R. Spolaczyk, and K. Merkel, "Digital wavefront measuring interferometry: Some systematic error sources," Applied Optics, Vol 22, Nov. 1983, pp. 3421–3432.

[10] B. Zhao and Y. Surrel, "Phase shifting: six–sample self–calibrating algorithm insensitive to the second harmonic in the fringe signal," Opt. Eng. Vol. 34, Sept., 1995, pp. 2821–2822.

[11] M. Milman, "An optimization approach to the suppression of vibration errors in phase shifting interferometry," JOSA A, May, 2002, pp. 992–1004.

[12] J. E. Grievenkamp, "Generalized data reduction for heterodyne interferometry," Opt. Eng., Vol. 23, July, 1984, pp. 350–352.

[13] M. Milman and S. Basinger, "Error sources for white light fringe estimation at low light levels," Applied Optics, May, 2002, pp. 1–17.

SIM White Light On–board Processing Algorithms

Mark H. Milman and Joe Catanzarite

Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109

1. Introduction Interferometry in optical astronomy is an important and growing field of astronomical observation. A number of stellar interferometers have come online over the past several years (e.g. Palomar [1], NPOI [2], REGAIN [3]), and several more are due to be operational in the near future (e.g. Keck [4], VLTI [5]). In addition, space based interferometers are also planned missions of NASA's Origins program [6], including the Space Interferometry Mission (SIM), the focus of the present paper. The fundamental measurement made by each of these interferometers is the white light fringe measurement to determine the optical pathlength delay between the two arms of the interferometer. SIM makes white light measurements with three independent interferometers observing three different objects. Two of these are the "guide" interferometers that observe bright objects (approximately 7th magnitude) to track the rigid body motion of the instrument. The third interferometer, the "science" interferometer, observes the science targets of interest.

A common method for making the white light measurement is to disperse the interfered starlight across multiple spectral channels and employ a phase shifting interferometry (PSI) algorithm to determine the phase difference in each channel. In the phase–delay method, knowledge of the mean wavenumber of each channel is used to convert the phase measurement to a delay measurement. The individual delays are then combined to produce a single delay. There are many alternatives for performing this last function – the delays can simply be averaged, or they can be combined in a weighted least squares sense, etc. When this delay measurement is combined with a metrology measurement of the pathlength difference of the starlight as it traverses the optical system, the fundamental observable of the instrument is constructed: the "external" pathlength delay, which is mathematically equivalent to the projection of the interferometer baseline vector onto the star direction vector. From these measurements the astrometric parameters of the star can be estimated.

This paper focuses on several of the important aspects of how this fundamental observable is constructed, and some of the errors that arise in its development. There is a significant history of PSI algorithms that deal with commonly associated errors in interferometry measurements. These include problems associated with small signal, changes in the pathlength (e.g. vibration) as the measurement is being made, non–monochromatic light with error in the mean wavelength in spectral channels, etc. In many instances algorithms can be developed to overcome these errrors ([7],[8], [9]) Here we will tie together the interferometric measurement with the metrology measurement especially to overcome the effects of vibration, and errors inherent in the modulating element. Also some preliminary investigation into the impact of non–monochromatic light has PSI algorithms will be made.

1

2. Background. SIM makes the pathlength delay measurement by a combination of internal metrology measurements to determine the distance the starlight travels through the two arms of the interferometer, and a measurement of the white light stellar fringe to find the point of equal pathlength. Figure 1 defines the basic geometry of the interferometer. Assume here that the optical axes of the two telescopes that comprise the interferometer apertures are aligned with the star position vector s. The planes P_X and P_Y are two planes of equal phase for the planar wavefront of the starlight, and each is normal to s. The light through the two arms combine at the beam combiner located at z. Let l_X and l_Y denote the internal optical pathlength through the X and Y arms, respectively, to z. And let T_X and T_Y denote the total pathlengths of the starlight through the X and Y arms. Define

$$E_X = T_X - l_X; \quad E_Y = T_Y - l_Y. \tag{1}$$

Then $E_Y - E_X$ is the external pathlength, and defines the basic astrometric equation which relates the star direction vector and the interferometer baseline vector,

$$E_Y - E_X = \langle s, Y - X \rangle, \tag{2}$$

where $Y - X$ is the baseline vector of the interferometer. Now $E_Y - E_X$ is not directly observable, but from the equation (deduced from (1))

$$E_Y - E_X = (T_Y - T_X) - (l_Y - l_X), \tag{3}$$

we see that each of the quantities on the right is observable. $T_Y - T_X$, the total pathlength difference, is measured by means of white light fringe estimation, and $l_Y - l_X$, the internal pathlength difference is measured by a metrology system employed by the instrument.

Eq. (2) is typically written as

$$d_{ext}(t) = \langle b(t), s \rangle + \nu(t) \tag{4}$$

where d_{ext} is the instantaneous external pathlength delay, $b(t)$ is the instantaneous interferometer baseline vector, and ν is the measurement error. The fundamental objective of the instrument is to make measurements so that the astrometric parameters of the star vector s can be determined. Additional "guide" interferometers and an external metrology system that ties together all of the interferometers are used to track the baseline vector $b(t)$. Because a finite integration time is required to make this measurement, the basic model for the astrometric equation has the form

$$\bar{d}_{ext} = \langle s, \bar{b} \rangle + \nu, \tag{5}$$

where the overbar represents a time–averaged quantity. The emphasis of the remainder of the paper is the synthesis of $\bar{d}_{ext}$ via time–averaging of the white light fringe estimates and the internal metrology measurements over the observation period. Obtaining $\bar{b}$ involves a similar process using the guide interferometers and auxillary metrology measurements to relate the three interferometer baseline vectors. This process, sometimes refered to as "baseline regularization" [*] will not be discussed here.

3. Measuring the external delay. As described in the previous section, the external delay measurement, d_{ext} is synthesized from two other delay measurements – the total delay and the

internal delay. Henceforth these latter two delays will be denoted d_{tot} and d_{int}. respectively. The three different delays are related:

$$d_{tot} = d_{ext} + d_{int}. \tag{6}$$

The initial focus is on how to compute the average external pathlength delay for a single white light fringe measurement that requires, say τ sec of integration time. We will assume that this integration time coincides with the modulation period for the phase shifiting interferometry (PSI) measurement. A nominal value for τ we will use corresponding to a typical science target is 100ms; although dim targets may require several seconds of integration time. We will first take a small digression to explain how d_{tot} is ideally measured using white light interferometry

A general perspective of building phase estimators uses the following simple idea. We start with the fundamental interferometric intensity equation for monochromatic light

$$y = I_0[1 + V(\cos(kx + \phi))], \tag{7}$$

where y is the observed intensity, I_0 is the dc intensity, V denotes the visibility, k is the wavenumber of the monochromatic light (or mean wavenumber over a spectral channel), x is a known modulator displacement, and ϕ is the sought after phase. Phase shifting interferometry entails introducing known pathlength changes via the x variable in the expression above to set up a nonlinear system of equations to solve for all of the unknown variables. The common way for solving this system is to introduce the state X consisting of three components, $X \equiv (I_0, I_0 V \cos(\phi), I_0 V \sin(\phi))$. It can be shown [7] that the system of equations is *linear* in X,

$$Y = AX, \tag{8}$$

where Y is the vector of observed photon counts at different values of pathlength change variable x, and A is a known matrix that maps the state vector into the observation vector.

An N–bin algorithm essentially uses N values of x in (7) to generate N values of the observed intensity y. Hence, N equations in the three unknown variables. For an $N-$ bin "integrating bucket" algorithm (see [12]) centered at zero we define s as the total stroke length of the phase modulator and we let λ denote the wavelength of the light. From these two variables we define $\gamma = s/\lambda$ and $\Delta = 2\pi\gamma/N$. In the integrating bucket method, the modulator linearly sweeps through $2\pi\gamma$ radians from $-\pi\gamma$ to $\pi\gamma$. The design matrix A for the integrating bucket method that relates photon counts to the state is obtained by integrating (7) over the "buckets" $(u_i - \Delta/2, u_i + \Delta/2)$ with respect to x. The resulting A matrix has the form (cf [12])

$$A = \begin{bmatrix} \Delta & \cos(u_1) & -\sin(u_1) \\ \vdots & \vdots & \vdots \\ \Delta & \cos(u_N) & -\sin(u_N) \end{bmatrix}, \tag{9}$$

where $u_i = (i - 1/2 - N/2)\Delta$. The estimate of X is obtained from *any* unbiased linear estimator of the state, and the phase is subsequently derived from X via $\phi = \arctan(X_3/X_2)$. Once the phase has been determined, the delay is calculated by using the wavelength of the light. To avoid multiple wavelength ambiguities in the delay, the white light is dispersed into multiple spectral channels, each with a different mean wavelength. The ambiguities can be resolved by using the multiple wavelengths [14]. A question that is addressed in Section 5 is the validity of using monochromatic algorithms in a spectral channel that has a nonvanishing width.

3

So keping in mind the need to modulate the pathlength in order to measure it, we write

$$d_{tot}(t) = d_{ext} + d_{int} + d_{pzt}(t), \tag{10}$$

where $d_{pzt}(t)$ is the pathlength introduced by the modulation. If everything is "perfect", the white light fringe estimate is an estimate of $d_{ext} + d_{int}$, which presumes that this value is constant over the τ second integration time. (In reality this is not the case, and this important non–ideality will be discussed in the next section.) Call this measurement y_{wl};

$$y_{wl} = d_{ext} + d_{int}. \tag{11}$$

Keeping in mind that the internal pathlength is changing due to the phase modulation, the measurement made by internal metrology has the form

$$m(t) = d_{int} + d_{pzt}(t). \tag{12}$$

The average value of m is d_{int} since $\bar{d}_{pzt} = 0$, thus

$$d_{ext} = y_{wl} - \bar{m}. \tag{13}$$

4. Mechanical white light error sources and some fixes. There are many different sources of errors contributing to white light fringe estimation. In this section we will focus on the class that is generated by changes in the pathlength while the phase is being measured. When the pathlength is changing, even if we interpret eq.(13) as the average external pathlength difference over the τ second integration period (which is indeed the quantity that is required for the astrometric relationship in (2)), there in general will be an error because y_{wl} may not compute the average total pathlength delay over the integration time.

One immediate way of generating an error of this type is if the motion of the modulator deviates from the motion assumed in the phase estimation algorithm, i.e. the matrix A in (9) which is constructed from an ideal modulator motion is in error. A straightforward way around this problem is to use the internal metrology measurement to create a new matrix A with every modulation stroke. Unfortunately this is not entirely desirable because in addition to building this matrix for each phase measurement, its pseudoinverse also has to be constructed to generate the phase delay. An alternative to this is to use an algorithm based on an assumed and fixed trajectory of the modulating element, and then make corrections to the computed phase based on measured values obtain from internal metrology. The main advantage here is simplicity, with payoffs including a reduced computational load, and very importantly, a simpler analysis of errors.

Following this tack, denote the modeled trajectory by d_{pzt}^{mod} and write

$$d_{pzt}(t) = d_{pzt}^{mod}(t) + [d_{pzt}(t) - d_{pzt}^{mod}(t)]. \tag{14}$$

Upon defining

$$d_{int}^* = d_{int} + [d_{pzt}(t) - d_{pzt}^{mod}(t)], \tag{15}$$

(10) becomes

$$d_{tot}(t) = d_{ext} + d_{int}^*(t) + d_{pzt}^{mod}(t). \tag{16}$$

4

So now

$$y_{wl} = d_{ext} + \bar{d}_{int}^*, \qquad (17)$$

with metrology measurements

$$m(t) = d_{int}^* + d_{pzt}^{mod}(t). \qquad (18)$$

Averaging the metrology measurements yields

$$\bar{m} = \bar{d}_{int}^*, \qquad (19)$$

since $\bar{d}_{int}^{mod} = 0$. Thus

$$y_{wl} - \bar{m} = \bar{d}_{ext}, \qquad (20)$$

so that

$$\bar{d}_{ext} = y_{wl} - \bar{m}. \qquad (21)$$

Although we have apparently circumvented the need for redefining an A matrix with each stroke, the probelm is not completely solved. We have introduce the time varying term $d_{pzt}(t) - d_{pzt}^{mod}(t)$ into the total path.

2. White light systematic errors. Define $\delta_{ext}(t)$ and $\delta_{int}^*(t)$ by

$$\delta_{ext} = d_{ext}(t) - \bar{d}_{ext}, \quad \delta_{int}^*(t) = d_{int}^*(t) - \bar{d}_{int}^* \qquad (22)$$

and set

$$\delta \equiv \delta_{ext} + \delta_{int}^*. \qquad (23)$$

Suppose an $N-$bin integrating bucket algorithm is used to convert the vector of photon counts $(y_1, ..., y_N)$ into phasor estimates $I_0 V \cos(\phi), I_0 V \sin(\phi)$ by the $3 \times N$ gain matrix $K = (k_{ij})$,

$$I_0 V \cos(\phi) = \sum_{j=1}^{N} k_{2j} y_j, \quad I_0 V \sin(\phi) = \sum_{j=1}^{N} k_{3j} y_j, \qquad (24)$$

where $I_0 =$ intensity, $V =$ visibility and $\phi =$ phase. The true average delay over the single phase measurement period is $\bar{d}_{ext} + \bar{d}_{int}^*$ with associated phase ψ,

$$\psi = \frac{2\pi}{\lambda}(\bar{d}_{ext} + \bar{d}_{int}^*), \qquad (25)$$

where λ denotes the wavelength of the light. In [9] it is shown that the error in the estimate using the phase derived from (24) is

$$\psi - \phi = \frac{1}{2\sin(\Delta/2)}\{\sum k_{3j}[\cos^2(\psi)\int_{u_i-\Delta/2}^{u_i+\Delta/2} \sin(u)\delta(u/2\pi\gamma)du + \sin(2\psi)/2 \int_{u_i-\Delta/2}^{u_i+\Delta/2} \cos(u)\delta(u/2\pi\gamma)du] +$$

$$\sum k_{2j}[\sin(2\psi)/2 \int_{u_i-\Delta/2}^{u_i+\Delta/2} \sin(u)\delta(u/2\pi\gamma)du + \sin^2(\psi)\int_{u_i-\Delta/2}^{u_i+\Delta/2} \cos(u)\delta(u/2\pi\gamma)du]\} \qquad (26)$$

Note that in general the white light error due to the pathlength variation δ is also a function of the offset ψ. For SIM applications ψ is typically small (a fraction of a radian), and the first term dominates the error.

5

A frequency response of the error can be generated for any given gain matrix K by fixing a phase offset for ψ and computing the resulting error for deviations of the form $\delta(t) = \sin(\omega t) - \mu(\omega)$, where $\omega = 1$ corresponds to the modulation frequency and $\mu(\omega)$ is the average value of $\sin(\omega t)$ over the modulation interval. Because the astrometric observable entails an average value of the external pathlength delay over several seconds of integration time on a typical science object, an important quantity is the average value of the error due to sinusoid variation of frequency ω over many, say M, modulation periods. (For example, if a 30sec integration time is required and a delay measurement is made every .1sec, then $M = 300$.) In [10] this error is shown to have a form (ignoring the variations in ψ from measurement to measurement)

$$E_{ave}(\omega) = f(\omega)C_M(\omega) + g(\omega)S_M(\omega), \tag{27}$$

where

$$C_M(\omega) = \frac{1}{M} \sum_{k=0}^{M-1} \cos(2\pi k\omega), \quad S_M(\omega) = \frac{1}{M} \sum_{k=0}^{M-1} \sin(2\pi k\omega), \tag{28}$$

and the functions f and g decrease as $1/\omega$ with increasing frequency. The sums above can be computed analytically to obtain (using the complex exponential forms for sin and cos and recognizing the sums as a geometric series),

$$S_M(\omega) = \frac{\sin(2(M-1)\pi\omega) + \sin(2\pi\omega) - \sin(2M\pi\omega)}{2M(1 - \cos(2\pi\omega))}, \tag{29}$$

and

$$C_M(\omega) = \frac{1 - \cos(2\pi\omega) + \cos(2\pi(M-1)\omega) - \cos(2M\pi\omega)}{2M(1 - \cos(2\pi\omega))}. \tag{30}$$

Thus when ω is an integer, there is no attenuation in the error due to averaging. But when it is not, the error will decrease as $1/M$ for each frequency ω. Because of the denominator terms, it is seen that the attenuation is slower when the frequency is close to an integer. The oscillation in the error will also be slower for frequency values near an integer. This oscillation frequency is $\omega_0 \equiv \omega \bmod 1$, so that with M measurements there would be $M\omega_0$ total oscillations. In contrast to the case of a periodic disturbance with an integer frequency, the expectation for an arbitrary signal based on Fourier analysis is that the average error diminishes with time (increasing number of intervals). Thus, periodic disturbances that are close to an integral multiple of the modulator frequency will attenuate slowly (and not at all if the disturbance is an exact multiple).

The error formula (26) can be used to correct the phase error if the deviation function δ is known. We will begin by assuming this is the case, and then discuss how this can be implemented and identify the residual errors.

Simplifying the notation, we may write (26) as

$$\psi - \phi = \int G(u, \psi)\delta(u)du, \tag{31}$$

where $G(u, \psi)$ is formed directly from the gain definition and the true phase ψ. Suppose $\delta(u)$ is known, and write $e = \psi - \phi$. Then (31) can be written as

$$e = \int G(u, \phi)\delta(u)du + \int \frac{\partial G(u, \phi)}{\partial \phi} e\delta(u)du. \tag{32}$$

6

Then,

$$e = \int G(u, \phi)\delta(u)du(1 - \int \frac{\partial G(u, \phi)}{\partial \phi}\delta(u)du). \tag{33}$$

As $\delta(u)$ is presumably small, the second term in the expression above should be ignorable, and we are left with

$$\psi = \phi + \int G(u, \phi)\delta(u)du. \tag{34}$$

The implementation of the correction term in (34) is straightforward. Let (k_{ij}) denote the $3 \times N$ components of the gain matrix. Then for $i = 1, ..., N$ the following quadratures are computed:

$$s_i = \int_{u_i - \Delta/2}^{u_i + \Delta/2} \sin(u)\delta(u/2\pi\gamma)du, \qquad c_i = \int_{u_i - \Delta/2}^{u_i + \Delta/2} \cos(u)\delta(u/2\pi\gamma)du. \tag{35}$$

Then (36) is implemented as

$$\psi = \phi + \frac{1}{2\sin(\Delta/2)}\{\cos^2(\psi)\sum_{i=1}^{N} k_{3i}s_i + \frac{\sin(2\psi)}{2}\sum_{i=1}^{N} k_{3i}c_i + \frac{\sin(2\psi)}{2}\sum_{i=1}^{N} k_{2i}s_i + \sin^2(\psi)\sum_{i=1}^{N} k_{2i}c_i\}. \tag{36}$$

Recall that δ is the deviation about the mean of the sum of the external path and the modified internal path. This quantity is actually somewhat observable for the science interferometer.

4. Sources of error. The mechanical sources for error in the determination of the white light estimate are those mechanisms that directly contribute variations about the mean of the external path and variations about the mean of the modified internal path defined in (1) and (15), respectively.

Since the star direction vector may assumed to be fixed in inertial space over the duration of the observation, the only way the external path can change is by a change in the inertial position of the interferometer baseline vector. This vector is defined as the difference between the two fiducial positions on the siderostat mirrors, and is thus affected by both rigid body motion of the instrument and vibrations propagating through the siderostats. The only means for active control of the external pathlength is via the attitude control system. Because the bandwidth of the attitude control system is between .1Hz and .01Hz, and the modulation period is typically less than a second (.1sec for grid star, faster for a guide star), the rigid body contribution to δ_{ext}^* is the dominant term for the science interferometer and is essentially linear. (A representative disturbance spectrum for the felxible body contribution is being developed.) The pathlength control system compensates for the rigid body motion to maintain the total pathlength error to the 10nm requirement to ensure good fringe visibility. The controller therefore introduces a large change in the internal path for the science interferometer. Assuming the rigid body motion of the instrument is a sinusoid with amplitude of 10^{-5}rad (about 2asec) and a frequency of .01Hz, the maximum external pathlength velocity is about 6μm/sec. If a phase measurement is made every .1sec, as is assumed for the science interferometer, the pathlength controller must vary the internal pathlength by .6μm to compensate for the change in the external pathlength over the phase measurement time period to achieve the 10nm requirement. Hence, the presence of pathlength control mitigates the external and internal pathlength variations that contribute to the white light estimation error. In fact the guide interferometers and external metrology provide the signal that drives the compensation and if the control system acts ideally, the internal pathlength changes introduced by the controller completely cancel the external pathlength changes.

7

In the real situation the compensation is effective over some specified bandwidth, and a residual variation remains with spectral content beyond the bandwidth of the controller. Also what remains uncompensated are the inaccuracies of the signal, the implementation of the controller, internal vibrations/drifts, and modulator motion error. Using a fixed gain phase measurement algorithm, the contribution of modulator motion to the error is the difference between the assumed motion and the true motion of the modulator.

5. Non—monochromatic problem. The analysis thus far has centered around the monochromatic light case. Modifications are necessary when the spectral bandwidth is not vanishingly small. In this section we will take a preliminary look at some of the parameters that may mandate these modifications. We examine three parameters: the number of spectral channels (which determines the channel width), the number of time bins, and the pathlength delay that is to be estimated. For example, increasing the number of channels eventually reduces to the monochromatic case. So one question is how many channels are necessary to make the problem look monochromatic? And when this is not possible, what modifications of the baseline monochromatic estimator are necessary to recover performance? We will be treating only the simplest model that assumes a rectangular bandpass with constant intensity/visibility within the channel. For simplicity we will only analyze the situation using a phase stepping (as opposed to integrating bucket) modulation scheme.

When the monochromatic estimator is not sufficient for this model there are essentially two modifications that can be made. The simplest one involves including a known sinc function in the matrix equation that relates the state variables (intensity and phasor quantities) to the measured intensities. The more complex model contains the unknown delay as part of the sinc term. This model leads to a nonlinear estimation problem to determine the state variables. We will focus attention throughout this discussion on a single spectral channel with mean wave number k_0 ($k_0 = 2\pi/\lambda_0$, with λ_0 denoting the corresponding wavelength) and channel width Δk. We will assume throughout that $\Delta k = 2\pi/(10^3 M)$ (units of 1/nm) where M denotes the number of spectral channels. (This value of Δk approximately corresponds to a spectral band between 500nm and 1000nm divided into M channels.) For a pathlength difference, x, the intensity model using a rectangular bandpass is [1]

$$I = I_0[1 + V\text{sinc}(\Delta kx/2)\cos(k_0x)]. \tag{37}$$

If δ denotes the unknown delay, and the pathlength is modulated in N equidistant steps, $x_i = i\lambda_0/N$, $i = 1, ..., N$, the estimation problem is to determine δ from the N equations

$$y_i = I_0[1 + V\text{sinc}(\Delta kx_i/2 + \Delta k\delta/2)\cos(k_0x_i + k_0\delta)] \tag{38}$$

In what follows y_i, k_0, Δk, and x_i are all assumed to be known, i.e., we are postulating a phase stepping method in which there is no measurement error, the wavenumbers and spectral channels widths have been precisely calibrated, and the phase steps have no error.

It is useful to relate the value of the sinc function in (1) to unity to capture the comparison between the models (2) and (3). If M channels are used and the modulation sweeps a micron, the maximum deviation of $\text{sinc}(\Delta kx/2)$ from unity is about $\pi^2/(6M^2)$; which provides some measure of the deviation between the two models. To truly quantify the errors in the estimate of the delay it is necessary to solve for the delays in the two models.

The ususal manipulation of (3) leads to the linear estimation problem from the system

$$
\begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} 1 & \cos(u_1) & -\sin(u_1) \\ \vdots & \vdots & \vdots \\ 1 & \cos(u_N) & -\sin(u_N) \end{bmatrix} \begin{bmatrix} I_0 \\ I_0 \cos(\phi) \\ I_0 \sin(\phi) \end{bmatrix}, \tag{39}
$$

where $u_i = k_0 x_i$ and $\phi = k_0 \delta$. (The slightly different form of the equations is due to the phase–stepping versus integrating–bucket modes of phase modulation.) We will write the nominal monochromatic system of equations as

$$
y = A_0 x. \tag{40}
$$

In the figure below we plot the true delay versus the error in the delay estimate based on the model (3). These errors were computed for four different combinations of time bins and spectral channels. The wavelength and modulation length are both 750 nm in these plots.
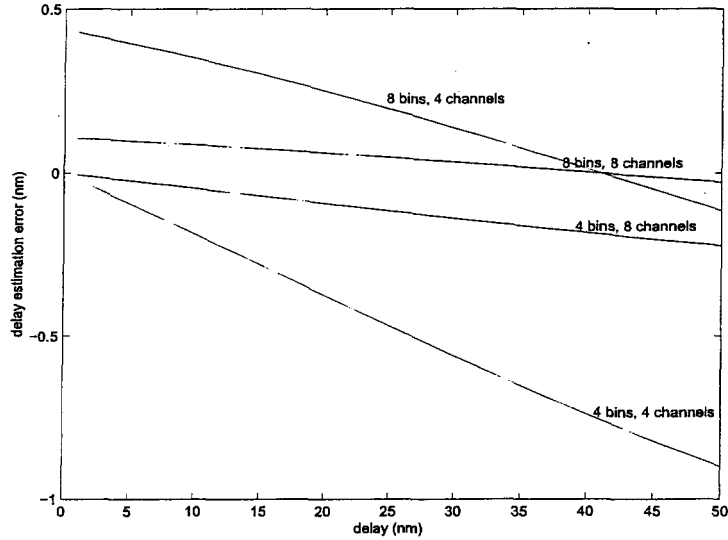

Figure 1. Phase estimation error using monochromatic alogrithm

Thus it is seen that using eight spectral channels produces a relatively small error (although not necessarily acceptable for all applications). This error is what is obtainable without modifying the basic algorithm.

9

In the next figure the number of channels is increased to 16, and two modulation lengths were used: 750 nm and 1000 nm, while the wavelength was fixed at 750 nm. The resulting error in the worst case is less than 50pm.
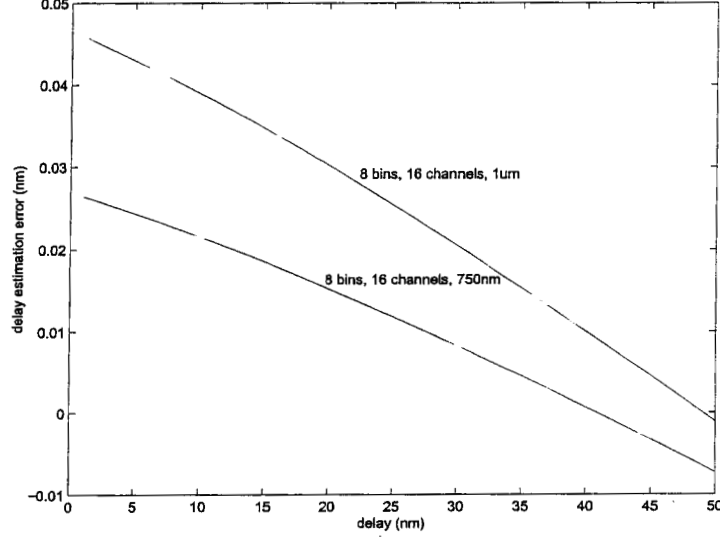

Figure 2. Phase estimation error using monochromatic alogrithm with 16 spectral channels

Next we model (1) using the sinc function, but only at the known modulated delays. This model has the form

$$
\begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} 1 & \mathrm{sinc}(\Delta k x_1/2)\cos(u_1) & -\mathrm{sinc}(\Delta k x_1/2)\sin(u_1) \\ \vdots & \vdots & \vdots \\ 1 & \mathrm{sinc}(\Delta k x_N/2)\cos(u_N) & -\mathrm{sinc}(\Delta k x_N/2)\sin(u_N) \end{bmatrix} \begin{bmatrix} I_0 \\ I_0\cos(\phi) \\ I_0\sin(\phi) \end{bmatrix}, \tag{41}
$$

which we write in matrix form as

$$
y = Ax. \tag{42}
$$

In anticipation of a larger nonlinearity that results when the number of spectral channels is decreased and/or the delay offset becomes larger, we introduce the matrix function $A(\delta)$,

$$
\begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} 1 & \mathrm{sinc}(\Delta k[x_1+\delta]/2)\cos(u_1) & -\mathrm{sinc}(\Delta k[x_1+\delta]/2)\sin(u_1) \\ \vdots & \vdots & \vdots \\ 1 & \mathrm{sinc}(\Delta k[x_N+\delta]/2)\cos(u_N) & -\mathrm{sinc}(\Delta k[x_N+\delta]/2)\sin(u_N) \end{bmatrix} \begin{bmatrix} I_0 \\ I_0\cos(\phi) \\ I_0\sin(\phi) \end{bmatrix}. \tag{43}
$$

With this notation, $A = A(0)$.

The phase estimate is recovered exactly as before, viz. $\hat{\phi} = \arctan(\hat{x}_3/\hat{x}_2)$ where $\hat{x} = A^\dagger y$. ($A^\dagger$ is the pseudoinverse of A.) It is important to realize that this is still a linear estimation problem. Figure 3 contains results, analogous to those in Figure 1. It is seen that an eight bin/eight spectral channel algorithm yields a very small error that should suffice for almost all applications. However,

10

the use of eight spectral channels may increase the read noise penalty to be unusable on dim stars. In the worst case (four bins, four channels, 50nm offset), an error of about 85 pm results.

Because the results are somewhat insensitive to the small error in the model of the sinc function (using $A(0)$ instead of $A(\delta)$), we conjecture that error analysis of this algorithm does not have to include the effects of (small) errors introduced through the sinc term in (1). Hence, much of the analysis should very closely follow the analysis that has been worked through for monochromatic light.
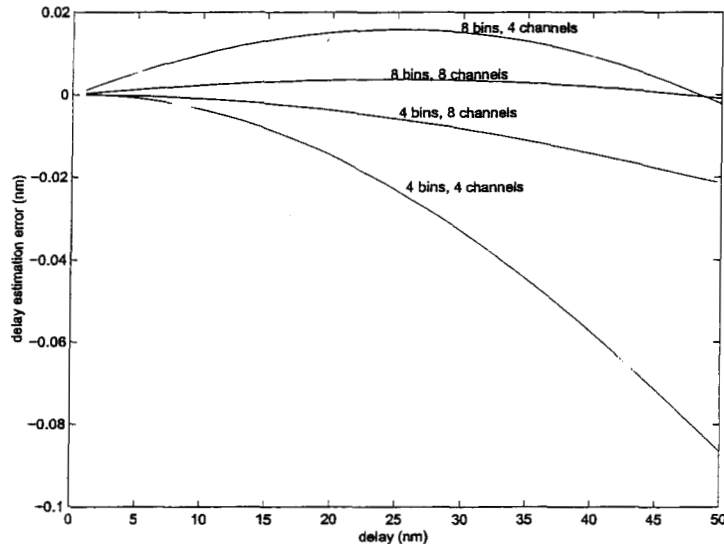

Figure 3. Phase estimation errors due to imprecise modeling for 4 configurations.

5.1. A nonlinear estimation scheme. For larger values of Δk and δ (or more stringent requirements) a nonlinear approach to the estimation problem is necessary. Thus we introduce the function

$$G(x) = [A(\delta(x))]^{\dagger} y, \tag{44}$$

where $\delta(x) = \tan^{-1}(x_3/x_2)$, and determine a fixed point of the map. That is we solve the equation

$$x = G(x). \tag{45}$$

This equation is solved via the fixed–point teration scheme

$$x_{k+1} = G(x_k), \quad x_0 = G(0) = A(0)^{\dagger} y. \tag{46}$$

Convergence of the scheme is motivated as follows. First we note that a fixed point exists. (This is guaranteed by (2). In the case of noise on the measurements, a regularity argument must be used to prove existence.) Given that a fixed point exists, local convergence of the iteration (11) is established if it can be shown that $|G'(x)| < 1$ in a neighborhood of the solution. An important approximation leading to this result is (see [3])

$$|A^{\dagger} - B^{\dagger}| \leq |A^{\dagger}||A - B|_F, \tag{47}$$

11

[3] D. Mourard, D. Bonneau, A. Glentzlin, G. Merlin, R. Pétrov, M. Pierron, N. Thureau, L. Abe, P. Bério, A. Blazit, O. Chesneau, P. Stee, S. Ragland, F. Vakili, and C. Verinaud, "The G12T/REGAIN Interferometer," Ibid, pp. 434–443.

[4] M. M. Colavita and P.L. Wizinowich, "Keck Interferometer: Progress report," Ibid, pp. 310–320.

[5] A. Glindemann, R. Abuter, F. Carbognani, F. Delplancke, F. Derie, A. Gennai, P. Gitton, P. Kervella, B. Koehler, S. Lévêque, S. Menardi, A. Michel, F. Paresce, T. P. Duc, A. Richichi, M. Schöller, M. Tarenghi, A. Wallander, and R. Wilhelm, "The VLTI interferometer: A unique instrument for high–resolution astronomy," Ibid, pp. 2–12.

[6] R. Danner and S. Unwin, *Space Interferometry Mission, Taking the Measure of the Universe*, JPL Pub. 400–811, March, 1999.

[7] K. Creath, "Phase–measurement interferometery techniques," in *Progress in Optics XXVI*, Elsevier Science Pub. B. V., 1988, pp. 350–393.

[8] P. J. de Groot, "Derivation of algorithms for phase–shifting interferometry using the concept of a data–sampling window," Appl. Opt., Vol. 34, August, 1995, pp. 4723–4730.

[9] J. Schwider, R. Burrow, K.–E. Elsner, J. Grzana, R. Spolaczyk, and K. Merkel, "Digital wavefront measuring interferometry: Some systematic error sources," Applied Optics, Vol 22, Nov. 1983, pp. 3421–3432.

[10] B. Zhao and Y. Surrel, "Phase shifting: six–sample self–calibrating algorithm insensitive to the second harmonic in the fringe signal," Opt. Eng. Vol. 34, Sept., 1995, pp. 2821–2822.

[11] M. Milman, "An optimization approach to the suppression of vibration errors in phase shifting interferometry," JOSA A, May, 2002, pp. 992–1004.

[12] J. E. Grievenkamp, "Generalized data reduction for heterodyne interferometry," Opt. Eng., Vol. 23, July, 1984, pp. 350–352.

[13] M. Milman and S. Basinger, "Error sources for white light fringe estimation at low light levels," Applied Optics, May, 2002, pp. 1–17.